

# Über die Hörbarkeit einzelner Teiltöne in Melodien aus harmonisch-komplexen Klängen

UWE BAUMANN

Fachgebiet Akustische Kommunikation, Technische Universität München

## Einleitung

Die Wahrnehmbarkeit einer Melodie in mehrstimmiger Musik hängt in starkem Maße von der Fähigkeit des Zuhörers ab, simultan auftretende Spektralkomponenten unterschiedlichen akustischen Objekten zuzuordnen und diese getrennt wahrzunehmen. Fig. 1 gibt eine Übersicht über Möglichkeiten der Akzentuierung und damit Heraushebung eines einzelnen Teiltones in einem harmonisch-komplexen Klang.

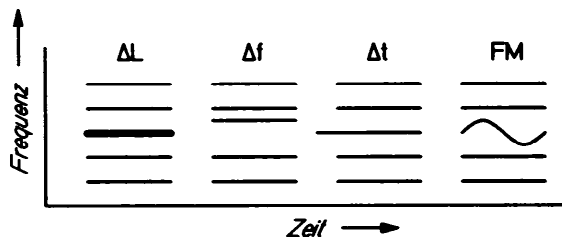


Fig. 1: Akzentuierung einer Spektralkomponente innerhalb eines harmonischen Klanges: Erhöhung des Pegels, Verstimmung einer Harmonischen, Asynchronität, Frequenzmodulation.

Abhängig vom Ausmaß der Beeinflussung von Pegel, Frequenz und Einsatzzeit einer oder mehrerer Spektralkomponenten ist eine getrennte Wahrnehmung der auf diese Weise akzentuierten Schallanteile neben dem harmonisch komplexen Grundklang möglich. Die Entwicklung eines Computermodells zur Nachbildung dieser Trennungsfähigkeit [1] erfordert die Kenntnis von Grenzwerten. Innerhalb dieser Grenzwerte tritt in der Regel die subjektive Fusion simultaner Spektralkomponenten ein, außerhalb ist das getrennte Heraushören von Schallanteilen möglich.

Die weitaus meisten Experimente untersuchten bisher Wahrnehmungsschwellen für Klangveränderungen oder fragten nur sehr pauschal nach der Trennbarkeit von Tonkomplexen. Die Frage nach den Kriterien, welche das Erkennen einer Melodiekontur im mehrstimmigen Zusammenspiel beeinflussen, bleibt dabei weitgehend unberücksichtigt.

Des Weiteren wird immer wieder die Vermutung diskutiert, wonach Frequenzmodulation (FM) mittels gesonderter Gestaltbildungsmechanismen die Fusion (oder andererseits die Separierbarkeit) von Schallen fördert [2]. Motiviert durch die Eigenschaft natürlicher Instrumente bei Einsatz des musikalischen Gestaltungsmittel „Vibrato“ alle Spektralkomponenten gleichsinnig, also frequenzkohärent zu beeinflussen, nimmt man insbesondere an, daß kohärente FM gegenüber inkohärenter die Fusion von Spektralkomponenten fördert und damit die Erkennung von musikalischen Stimmen erleichtert. Aus diesem Grund wurde ebenfalls der Einfluß von FM auf die Erkennbarkeit von Melodielinien untersucht.

## Methode

Zwei kleine Melodien, welche aus jeweils sechs harmonisch-komplexen Klängen bestanden (jeder Klang setzte sich aus zwanzig Harmonischen zusammen; Zeitdauer eines Klanges 0,2 Sekunden; Pause zwischen zwei aufeinanderfolgenden Tönen ebenfalls 0,2 Sekunden; 10 Millisekunden Ein- und Ausschalttrampe) wurden als „Vordergrund“ Melodiefigur verwendet. Abhängig von der

Versuchsreihe wurden in jedem Klang eine oder mehrere Harmonische verändert, so daß sich, abhängig vom Ausmaß der Beeinflussung, unter Umständen die Wahrnehmung einer zweiten (Hintergrund-) Melodie ergab.

Um Lerneffekte zu vermeiden und zusätzlich eine abwechslungsreiche Schalldarbietung zu erreichen, wurden innerhalb der Harmonischen der beiden Vordergrundmelodien vier verschiedene Hintergrundmelodien erzeugt. Diese Hintergrundmelodien erhielten die Verläufe AUF, AB, AUF/AB oder AB/AUF (Tab. I). Fig. 2 zeigt schematisch eine innerhalb der Vordergrundmelodie „A“ durch FM induzierte, absteigende Hintergrundmelodie.

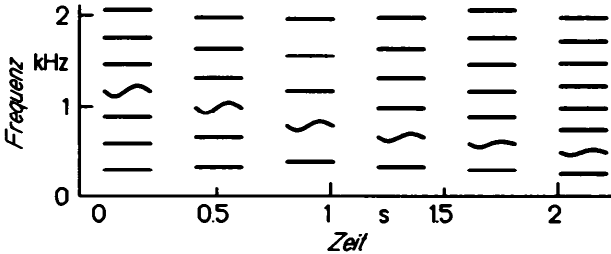


Fig. 2: FM induzierte Akzentuierung einer absteigenden Hintergrundmelodie (Melodie „A“ Vordergrund).

		1	2	3	4	5	6	Ton-Nummer	
A	Grundfreq.	293	329	392	329	293	246	Vordergrund	
	Tonbes.	D4	E4	G4	E4	D4	H3		
B	Grundfreq.	164	220	246	261	329	293		
	Tonbes.	E3	A3	H3	C4	E4	D4		
A	AUF	2	2	2	3	4	5		Hintergrund
		D5	E5	G5	H5	D6	D#6		
	AB	4	3	2	2	2	2		
		D6	B5	G5	E5	D5	H4		
	AUF/AB	2	2	2	2	2	2		
		D5	E5	G5	E5	D5	H4		
	AB/AUF	4	3	2	3	4	-		
		D6	H5	G5	H5	D6	D#6		
B	AUF	2	2	2	2	2	3	Hintergrund	
		E4	A4	H4	C5	E5	A5		
	AB	7	5	4	3	2	2		
		D6	C#6	H5	G5	E5	D5		
	AUF/AB	5	4	4	3	2	2		
		G#5	A5	H5	G5	E5	D5		
	AB/AUF	5	3	2	2	2	3		
		G#5	E5	H4	C5	E5	A5		

Tab. I: Die obere Tabelle listet Grundfrequenz und korrespondierenden Tonnamen für die sechs Töne der Melodien „A“ und „B“ auf. Die untere Tabelle zeigt den Index der modifizierten Harmonischen (oben) und den entsprechenden Tonnamen (unten) für jede Hintergrundmelodie.

Drei Versuchsreihen wurden durchgeführt. In jedem Versuch wurde eine Reihe von Akzentuierungsmöglichkeiten untersucht. In Versuchsreihe 1 wurde in jedem Klang eine Harmonische modifiziert. Dabei wurden vier Möglichkeiten untersucht: die Modifizierung des Einsatzzeitpunktes  $\Delta t$  (gestuft von -10 bis 30 Millisekunden)<sup>1</sup>; die Beeinflussung des Pegels  $\Delta L$  (-10 bis 10 dB); die Veränderung der relativen Frequenz einer Harmonischen  $\Delta f$  (-2% bis 2%); die Änderung der relativen Frequenzmodulationstiefe  $M$  (0,5% bis 6%, Modulationsrate 5 Hz). In Versuchsreihe 2 wurden drei im Frequenz-Verhältnis 1:2:4 stehende Teiltöne gemeinsam verändert. Es wurden die gleichen Parameter wie in Versuch 1 variiert. Um den Einfluß von kohärenter FM auf die Erkennungsrate der Hintergrundmelodie zu untersuchen, wurde neben kohärenter FM ( $M_c$ ) auch eine nicht kohärente ( $M_{nc}$ ) Modulation der drei Teiltöne vorgenommen. Die  $M_{nc}$  Bedingung wurde durch Veränderung der Startphase des Modulators ( $\varphi_{M_i}$  in  $\pi$ ,  $3/2\pi$  und 0) erreicht. Um die Auswirkungen eines natürlicheren Klanges auf die Erkennungsrate der Hinter-

<sup>1</sup>negatives Vorzeichen äquivalent zum Beginn der modifizierten Harmonischen vor dem Rest.

grundmelodie zu ermitteln, wurde in Versuchsreihe 3 der Klang einer synthetischen Trompete anstelle des statischen, harmonisch-komplexen Klanges aus Versuch 1 und 2 verwendet.

Die Aufgabe der Versuchspersonen bestand darin, den Verlauf der Hintergrundmelodie im Vergleich mit einer vorangehenden, unveränderten Darbietung anzugeben.

Sämtliche Versuchsschalle wurden vor Beginn des Versuches digital generiert. Die Versuchspersonen befanden sich in einer schallisolierten Meßkabinen; die Darbietung von Anker- und Testschall erfolgte diotisch in zufälliger Reihenfolge über einen freifeldentzerrten Kopfhörer (Abtastfrequenz 16kHz, 16 Bit D/A). Jede Hintergrundmelodie wurde für jeden Parameter zwanzigmal ohne Feedback dargeboten. Die Anzahl der Versuchspersonen variierte; an Versuch 1 und 2 nahmen fünf Personen, an Versuch 3 sieben Personen teil.

Abhängig vom Anteil der erkannten Hintergrundmelodien und dem Grad der Beeinflussung wurde für jede Akzentuierungsart eine psychometrische Funktion ermittelt. Der Grenzwert für die Erkennung einer vorgegebenen Hintergrundmelodie wurde dabei auf 62,5% richtiger Antworten festgelegt. In der Regel wurde zur Bestimmung dieses Wertes zwischen den gemessenen Stützwerten linear interpoliert (Fig. 3).

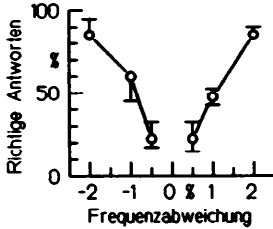


Fig. 3: Gemittelte Ergebnisse (Median und Interquartile) für Experiment 2, Parameter  $\Delta f$ . Die Ordinate gibt den Prozentsatz richtig erkannter Hintergrundmelodien an. Die Wahrscheinliche Schwankung wurde über die wiederholten Messungen und fünf Versuchspersonen ermittelt.

## Ergebnisse

In Tab. II sind die aus den psychometrischen Funktionen ermittelten Erkennungsschwellen für die verschiedenen Parameter aufgeführt.

		1 Harmonische veränd.				3 Harmonische veränd.				
Parameter		$\Delta t$	$\Delta L$	$\Delta f$	$M$	$\Delta t$	$\Delta L$	$\Delta f$	$M_c$	$M_{nc}$
Erkennungsschwelle	+	20 ms	4 dB	1.5%	1.5%	25ms	3 dB	1.5%	1.5%	1%
	-	4 ms	X	1%		4ms	X	1%		

		resynth. Trompete		
Parameter		$\Delta t$	$\Delta L$	$\Delta f$
Erkennungsschwelle	+	22 ms	3 dB	1.5%
	-	> 10 ms	X	1.5%

Tab. II: Erkennungsschwellen für die Identifizierung einer Hintergrundmelodie, die in den Harmonischen einer Vordergrundmelodie versteckt wurde. (X - nicht erkannt)

Die Versuchspersonen hatten generell keine Probleme mit der Aufgabenstellung. Auch Versuchspersonen ohne Erfahrung in Hörversuchen konnten die Versuchsreihen absolvieren. Fast alle Versuchspersonen gaben an, nach einer Anzahl von Darbietungen auch im unbeeinflussten Ankerschall eine Hintergrundmelodie wahrzunehmen.

## Diskussion

### Einsatz-Asynchronität ( $\Delta t$ )

Rasch zeigte in [3] den großen Einfluß von Einsatzzeitdifferenzen auf die Mithörschwellen von

quasi-simultanen Tonpaaren. Die Ergebnisse der Versuche 1 und 2 bestätigen diese Aussage mit dem Zusatz, daß sich ein asymmetrisches Verhalten in der Erkennungsschwelle der Hintergrundmelodie zeigt. Erfolgt der Einsatz der beeinflussten Teiltöne vor dem Beginn der Vordergrundmelodie, erkennen die Versuchspersonen schon ab einer Verzögerungszeit von  $-4$  Millisekunden den Verlauf der Hintergrundmelodie; bei einem Einsatz nach dem Beginn der Vordergrundmelodie wird die Hintergrundmelodie erst nach einer Verzögerung von  $20-25$  Millisekunden erkannt. Aus dem Vergleich mit den den Ergebnissen von Versuch 3 läßt sich entnehmen, daß Schalle mit dynamischem Einschwingverhalten die Erkennung von Hintergrundmelodien erst bei wesentlich längeren Verzögerungszeiten erlauben.

#### *Diskontinuität der spektralen Hüllkurve ( $\Delta L$ )*

Tab. II zeigt, daß sich lediglich durch eine Erhöhung des Pegels einzelner Teiltöne die Erkennung einer Hintergrundmelodie erreichen läßt. Eine Absenkung führt zwar zu deutlich wahrnehmbaren Klangveränderungen, die Erkennungsrate liegt aber innerhalb der Ratewahrscheinlichkeit. Der Vergleich mit den Ergebnissen von Versuch 3 zeigt, daß im Unterschied zu den Ergebnissen für die Einsatz-Asynchronität ein dynamisches Signal mit komplexer spektraler Hüllkurve die Erkennungsschwelle nicht beeinflusst.

#### *Statische Verstimmung von Harmonischen ( $\Delta f$ )*

Die Ergebnisse der Versuche zeigen, daß bereits ab einer relativen Verstimmung einzelner Harmonischer zwischen  $1\%$  und  $1,5\%$  die Erkennung einer Hintergrundmelodie möglich ist. Versuch 3 zeigt, daß der Einfluß einer komplexen spektralen Hüllkurve auf die Erkennungsrate (verglichen mit Frequenzunterschiedsschwellen) klein ist.

#### *Periodische Verstimmung von Harmonischen (FM)*

Die relative Frequenzmodulationstiefe muß einen Wert von  $1,5\%$  (Spitze—Spitze Frequenzauslenkung  $3\%$ ) überschreiten, um die Wahrnehmung einer Hintergrundmelodie zu ermöglichen. Der Vergleich mit den Versuchsergebnissen für die statischen Verstimmung legt die Vermutung nahe, daß durch eine Periodisierung der Frequenzänderung keine wesentliche Erniedrigung der Erkennungsschwelle erfolgt. Weiterhin ergaben sich keine Hinweise für eine Verbesserung der Erkennungsleistung durch kohärente FM: Versuch 2 zeigt für Kohärenz ( $M_c$ ) gegenüber Inkohärenz ( $M_{nc}$ ) keine Absenkung der Erkennungsschwelle. Das Gegenteil ist der Fall: Inkohärenz führt eher zur Wahrnehmung der Hintergrundfigur als Kohärenz. Ebenso zeigen sich zwischen den Ergebnissen von Versuch 1 (FM einer einzelnen Harmonischen) und Versuch 2, Bedingung  $M_c$ , keine Unterschiede. Diese Unterschiede wären aber zu erwarten, wenn auf kohärenter FM basierende Gruppierungsregeln für die subjektive Integration von Spektralkomponenten eine relevante Bedeutung hätten.

Diese Arbeit wurde von der Deutschen Forschungsgemeinschaft im Rahmen des SFB 204 „Gehör“ München gefördert.

## Literatur

- [1] Baumann, U. (1993). Anforderungen, Möglichkeiten und Grenzen der automatischen Analyse von Musik. In: *Fortschritte der Akustik* (DAGA 93, Frankfurt), DPG Verlag, Bad Honnef, S. 896–899.
- [2] McAdams, S. (1989). Segregation of concurrent sounds. I: Effects of frequency modulation coherence. *J. Acoust. Soc. Am.*, 86, 2148–2159.
- [3] Rasch, R. (1978). The perception of simultaneous notes such as in polyphonic music. *Acustica*, 40, 21–33.